

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-162439

(43)Date of publication of application : 06.06.2003

(51)Int.Cl.

G06F 12/00

G06F 3/06

(21)Application number : 2001-358320

(71)Applicant : HITACHI LTD

(22)Date of filing : 22.11.2001

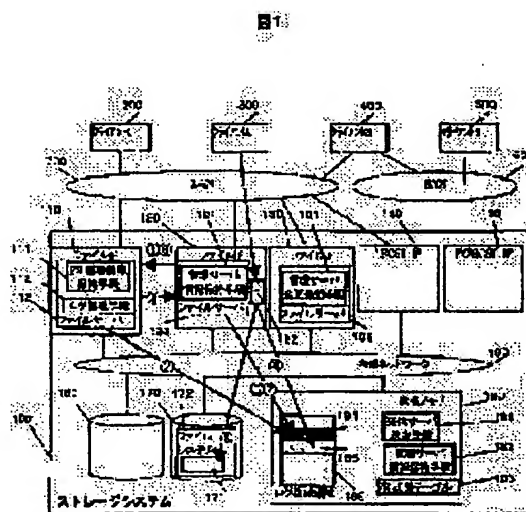
(72)Inventor : SONODA KOJI  
MATSUNAMI NAOTO  
KITAMURA MANABU  
TAKADA YUTAKA

## (54) STORAGE SYSTEM AND CONTROL METHOD THEREFOR

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide both interfaces of SAN and NAS for elimination of losing data, when troubles occur, and to enable accessing highly efficiently file systems which have arbitrary number of the same NAS interfaces.

**SOLUTION:** This system is provided with a plurality of interfaces for connecting to the outside, a plurality of disks 160, 170 which can be accessed by the plurality of interfaces, and a common memory 180 which can be accessed by the plurality of interfaces. The plurality of interfaces are block interfaces 140, 150 which process I/O requests on disk block units, and file interfaces mounting file servers 110, 120, 130 which process I/O requests on file units. In a part of the disks, a file system 172 which the file server shares with is constructed; and in the shared memory, a log storing domain which maintains the change logs of the file system and a management file-server information storing domain, which maintains information relating the management file-server which performs exclusive control of the file system and management of the log storing domain are constructed.



## LEGAL STATUS

[Date of request for examination]

18.07.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the  
examiner's decision of rejection or application  
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of  
rejection][Date of requesting appeal against examiner's decision  
of rejection]

[Date of extinction of right]



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-162439

(P2003-162439A)

(43) 公開日 平成15年6月6日 (2003.6.6)

(51) IntCl. <sup>7</sup>	識別記号	F I	ターミナル (参考)	
G 0 6 F 12/00	5 4 5	G 0 6 F 12/00	5 4 5 A	5 B 0 6 5
	5 3 1		5 3 1 J	5 B 0 8 2
	5 3 5		5 3 5 B	
3/06	3 0 1	3/06	3 0 1 A	
			3 0 1 Z	

審査請求 未請求 請求項の数11 O L (全 17 頁) 最終頁に続く

(21) 出願番号 特願2001-358320 (P2001-358320)

(22) 出願日 平成13年11月22日 (2001. 11. 22)

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 園田 浩二

神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム開発研究所内

(72) 発明者 松並 直人

神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム開発研究所内

(74) 代理人 100093492

弁理士 鈴木 市郎 (外1名)

最終頁に続く

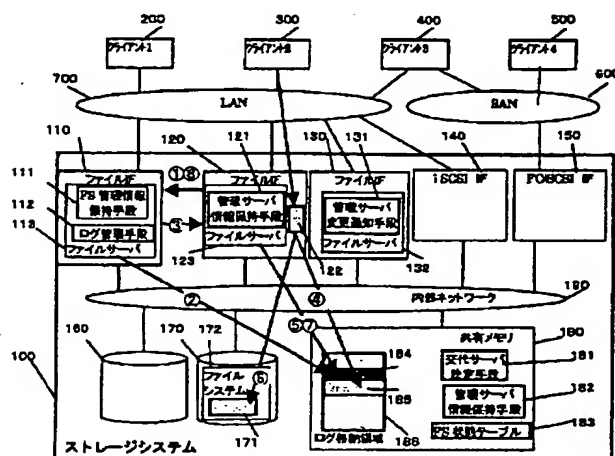
(54) 【発明の名称】 ストレージシステム及びその制御方法

(57) 【要約】

【課題】 SAN、とNASとの両インタフェースを提供し、障害発生時にもデータが失われることをなくし、また、任意の数のNASインタフェースが同一のファイルシステムに高性能アクセスすることを可能にする。

【解決手段】 外部接続用の複数のインタフェースと、複数のインタフェースからアクセス可能な複数のディスク160、170と、複数のインタフェースからアクセス可能な共有メモリ180とを備える。複数のインタフェースは、ディスクブロック単位のI/O要求を処理するブロックインタフェース140、150と、ファイル単位のI/O要求を処理するファイルサーバ110、120、130を搭載したファイルインタフェースとである。ディスクの一部には、ファイルサーバが共有するファイルシステム172が構築され、前記共有メモリには、前記ファイルシステムの変更ログを保持するログ格納領域と、前記ファイルシステムの排他制御及びログ格納領域の管理を行う管理用のファイルサーバに関連する情報を保持する管理ファイルサーバ情報格納領域とが構築されている。

図1



## 【特許請求の範囲】

【請求項1】 外部ネットワーク接続用の複数のインタフェースと、該複数のインタフェースからアクセス可能な複数のディスクと、前記複数のインタフェースからアクセス可能な共有メモリとを備えたストレージシステムにおいて、前記複数のインタフェースには、ディスクブロック単位のI/O要求を処理するブロックインターフェースと、ファイル単位のI/O要求を処理するファイルサーバを搭載したファイルインタフェースとのいずれかあるいは両方が搭載され、前記複数のディスクの一部には、複数のファイルサーバが共有アクセス可能なファイルシステムが構築され、前記共有メモリには、前記ファイルシステムの変更ログを保持するログ格納領域と、前記ファイルシステムの排他制御及びログ格納領域の管理を行う管理用のファイルサーバに関連する情報を保持する管理ファイルサーバ情報格納領域とが構築されていることを特徴とするストレージシステム。

【請求項2】 前記変更ログは、前記ファイルシステムの変更メタデータと、I/O要求に含まれる書き込みデータとの両方により構成されることを特徴とする請求項1記載のストレージシステム。

【請求項3】 前記管理ファイルサーバ情報は、前記ログ格納領域に前記変更メタデータのみを格納するか、変更メタデータと書き込み要求データとの両方を格納するかを示す設定情報であることを特徴とする請求項2記載のストレージシステム。

【請求項4】 前記ファイルシステムの管理ファイルサーバ以外のファイルサーバは、外部ネットワークから受信したファイルI/O要求がアクセスするファイルを格納しているファイルシステムの管理を行う管理ファイルサーバに対して、ファイル識別情報とアクセス領域情報とを含むファイルアクセス情報を送信し、その応答としてディスクブロック情報とログ格納アドレス情報とを受信する手段と、前記ファイルI/O要求に含まれる書き込みデータをログ格納アドレス情報に基づいて前記共有メモリに格納する手段と、前記書き込みデータを前記ディスクブロック情報に基づいて前記ファイルシステムが構築されているディスクに格納する手段とを備えることを特徴とする請求項2または3記載のストレージシステム。

【請求項5】 前記管理ファイルサーバは、他のファイルサーバからファイルアクセス情報を受信する手段と、前記受信したファイルアクセス情報を用いて該当するファイルをロックし、ディスクブロックを割り当て、該当するディスクブロック情報を算出するファイルシステムの管理手段と、前記ファイルアクセス情報を用いてログ格納領域内のログ格納アドレスを割り当てるログ格納領域管理手段と、ディスクブロック情報及びログ格納アドレス情報を、ファイルアクセス情報を送信してきたファイルサーバに送信する手段とを備えることを特徴とする

請求項4記載のストレージシステム。

【請求項6】 前記管理ファイルサーバは、前記ログ格納領域のサイズを設定するインタフェースを備えることを特徴とする請求項1ないし5のうちのいずれか1記載のストレージシステム。

【請求項7】 前記ファイルシステムの管理ファイルサーバ以外のファイルサーバは、前記管理ファイルサーバの障害発生時、前記管理ファイルサーバが管理していたログ格納領域に格納されている変更ログを用いてファイルシステムを復旧する手段を備えることを特徴とする請求項1ないし6のうちのいずれか1記載のストレージシステム。

【請求項8】 外部ネットワーク接続用の複数のインタフェースと、該複数のインタフェースからアクセス可能な複数のディスクと、前記複数のインタフェースからアクセス可能な共有メモリとを備えたストレージシステムの制御方法において、

前記複数のインタフェースには、ディスクブロック単位のI/O要求を処理するブロックインターフェースと、ファイル単位のI/O要求を処理するファイルサーバを搭載したファイルインタフェースとのいずれかあるいは両方が搭載され、前記複数のディスクの一部には、複数のファイルサーバが共有アクセス可能なファイルシステムが構築され、前記共有メモリには、前記ファイルシステムの変更ログを保持するログ格納領域と、前記ファイルシステムの排他制御及びログ格納領域の管理を行う管理用のファイルサーバに関連する情報を保持する管理ファイルサーバ情報格納領域とが構築されており、

前記ファイルシステムの管理ファイルサーバ以外のファイルサーバは、外部ネットワークからファイル書き込み要求を受信して、そのファイル書き込み要求を解析して書き込み対象ファイルが含まれるファイルシステムの管理ファイルサーバを特定し、管理ファイルサーバに対してファイル書き込み情報を送信した後、その応答としてユーザデータを書き込むディスクブロック情報とログ格納領域内に割り当てられたログ格納アドレス情報とを受信し、受信したログ格納アドレス情報を用いてユーザデータ格納領域にユーザデータを格納した後、ログ格納領域内のログステータス情報を変更し、ディスクブロック情報に基づいて前記ディスクにユーザデータを格納した後、ログ格納領域内のログステータス情報を変更し、ファイル書き込み結果情報を前記ファイルシステムの管理ファイルサーバに送信した後、外部ネットワークを經由して受信したファイル書き込み要求の応答を、外部ネットワークに送信することを特徴とするストレージシステムの制御方法。

【請求項9】 前記管理ファイルサーバは、管理ファイルサーバ以外のファイルサーバからファイル書き込み情報を受信すると、書き込み対象のファイルをロックし、前記ディスク上にユーザデータを書き込むディスクブ

ックを割り当て、ログ格納領域内のログ格納アドレスを割り当て、該ログ格納領域にファイルシステム管理データの変更情報とログステータス情報とを格納した後、前記ファイル書き込み情報を送信してきた管理ファイルサーバ以外のファイルサーバに、割り当てたディスクブロック情報とログ格納領域情報とを送信し、前記ファイル書き込み情報を送信してきた管理ファイルサーバ以外のファイルサーバからファイル書き込み結果情報を受信すると、その書き込み対象ファイルのロックを解除することを特徴とする請求項8記載のストレージシステムの制御方法。

【請求項10】 前記管理ファイルサーバ以外のファイルサーバは、前記管理ファイルサーバに障害が発生したことを認識した場合、前記管理ファイルサーバ情報を参照して、前記管理ファイルサーバが管理していたログ格納領域を特定し、特定したログ格納領域に格納されている変更ログを順次参照し、この変更ログが未反映のファイルシステムに対し変更処理を反映させ、全ての変更ログの反映処理の完了後、前記管理ファイルサーバが管理していた全てのファイルシステムの排他制御とログ管理領域とを引き継ぎ、管理ファイルサーバが変更されたことを前記管理ファイルサーバ以外の他のファイルサーバに通知することを特徴とする請求項8または9記載のストレージシステムの制御方法。

【請求項11】 前記変更ログの反映処理は、変更処理がファイル書き込みの場合、前記ログステータス情報を参照し、ログステータス情報がユーザデータ未書き込み状態となっていれば、ファイルシステムへの変更反映処理は行わず、前記ログステータス情報がユーザデータディスク書き込み完了状態となっていれば、その変更ログに応じて、ファイルシステム上のファイルシステム管理データのみを変更し、前記ログステータス情報がユーザデータログ書き込み完了状態となっていれば、変更ログに含まれるユーザ書き込みデータをファイルシステムに反映した後、ファイルシステム上のファイルシステム管理データを変更する処理であることを特徴とする請求項10記載のストレージシステムの制御方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ストレージシステム及びその制御方法に係り、特に、NAS機能及びSAN機能を搭載したRストレージシステム及びその制御方法に関する。

【0002】

【従来の技術】 近年のインターネット技術の進歩に伴い、Webアプリケーション、ストリーミングアプリケーション、Eビジネスアプリケーション等が急激に普及し、これらのアプリケーションが必要とするデータ量も飛躍的に増加してきており、日常生活やビジネスが要求するストレージ容量は爆発的に増加している。そして、

IBM Redbooksの“Storage Networking Virtualization”によれば、ストレージコストは確実に低下しているが、データの管理コストが増大するという問題点が発生してきている。

05 【0003】 前述の問題点を解決する従来技術として、Storage Area Network (SAN) と呼ばれる技術とNetwork Attached Storage (NAS) と呼ばれる技術が知られている。前述した“Storage Networking Virtualization”によると、SANは、ストレージ専用的高速ネットワークをFibre Channel を用いて構築し、分散されたデータを集約することにより管理コストを低減することができるようにしたものである。この技術は、ストレージ専用のネットワークを使用することにより、Local Area Network (LAN) におけるネットワークトラフィックの影響を排除することができるため、高性能なI/O性能を実現することができる。但し、SANが提供する機能は、ディスクブロックレベルの入出力機能であるため、異なるホストやOS相互間でファイルレベルでデータ共有することが困難であるという側面を持つ。

10 【0004】 一方、<http://www.sun.com/Storage/white-papers/NAS.html>の“NETWORK STORAGE SOLUTIONS White Paper Network-Attached Storage”によれば、NASは、NFSやCIFS等のネットワークプロトコルを通じたプラットフォーム独立のストレージ共有手段を提供するファイルサーバである。NASは、SANと同様に、分散されたデータを集約して管理コストを低減することができる。また、NASは、ファイルサーバとして最適化されており、LANに直接接続され、ファイルレベルのアクセス手段を提供している。

15 【0005】 前述したように、SAN、NASは、共にデータを集約し、管理コストを低減するための技術であるが、提供するアクセス手段が異なるため、用途に応じた使い分けが必要となる。また、前述のように、データを集約し多くのホストで共有利用する場合、高信頼性と高可用性とが重要な要素となる。

20 【0006】 前述したように、SANは、ストレージ専用のネットワークであり、その信頼性は、SANに接続される個々のストレージに依存し、ストレージとしてRAID (Redundant Array of Inexpensive Disks) を用いることにより高い信頼性を提供することができる。また、RAIDが複数のインタフェースを提供することにより、あるインタフェースに障害が発生した場合にも、別のインタフェースを用いてサービスを継続することができ、高い可用性を提供することができる。

25 【0007】 一方、NASは、ファイルシステムを持つファイルサーバであり、NASの信頼性は、ファイルサーバの信頼性そのものとなる。ところが、ファイルサーバは、RAIDを用いるだけでは高い信頼性を提供することができない。文献“最前線UNIX (登録商標) のカーネル”、ユーレッシュ・ヴァハリアの9、12、5

節によれば、UNIX（登録商標）のファイルシステムは、性能向上のためにメモリ上にバッファキャッシュを設け、複数の書き込み処理を纏めて、一括したディスク書き込みを行うようにしている。このため、この技術は、システムクラッシュ時にディスクに書き込まれていないデータが失われてしまう。失われるデータは、ファイルデータそのものと、ファイルシステムの構造を記述しているメタデータの2つに分類することができる。メタデータの変更が失われた場合、ファイルシステムに矛盾が生じ、ファイルシステムを使用できなくなるという問題が生じる。

【0008】このような問題を解決する方法として、“最前線UNIX（登録商標）のカーネル”、ユーレッシュ・ヴァハリアの11.7節に述べられているメタデータロギングの技術と、“最前線UNIX（登録商標）のカーネル”、ユーレッシュ・ヴァハリアの11.5節に述べられているログ構造ファイルシステムの技術が知られている。

【0009】メタデータロギングは、ディスク上に固定的に設けられた領域に、メタデータの変更ログを常に書き込み、システムクラッシュ時にこのメタデータ変更ログを参照して、ディスクに反映されていないメタデータの変更を反映させることにより、ファイルシステムの矛盾を解決するという方法である。メタデータロギングは、このような方法により、ファイルシステムの矛盾を発生させることをなくすることができるが、ファイルデータが失われる可能性は相変わらず残っている。

【0010】ログ構造ファイルシステムは、ファイルシステムに対して行われる変更をメモリ上の巨大なログエントリに蓄積し、ファイルシステムの一貫性を保持した状態で常にディスク上のログの最後尾に書き込むことにより、メタデータとユーザデータとの両方を保証することを可能としたものである。但し、このファイルシステムは、ディスク上のログに書き込まれる前にシステムクラッシュが生じた場合、メモリ上のログエントリに保存されていたデータが失われてしまうという問題を生じる。

【0011】このような問題を解決する技術として、[http://www.netapp.com/tech\\_library/1004.html](http://www.netapp.com/tech_library/1004.html)に記述されるNetwork Appliance社の“Using NUMA Interconnectsto Implement Highly Available File Server Appliances”が知られている。この技術は、NAS専用に最適化されたもので、不揮発性メモリ（NVRAM）とRAIDディスクとを持ち、RAIDディスク上にログ構造ファイルシステムを構築したものである。そして、NVRAMには、ネットワーク経由で受信したNFSコマンドを全てロギングし、RAIDディスクには、一貫性を保った状態のログを格納することとしている。

【0012】前述の技術は、ログエントリをRAIDディスクに書き込む前にシステムがクラッシュした場合に

も、システム復旧後にNVRAM上のNFSコマンドログを用いて、ファイルシステム処理を再実行してファイルシステムを完全な状態に復旧することができ、また、データを完全に保証することができる。

【0013】また、この技術は、ファイルサーバを搭載した2つのノードを独立したネットワークにより接続し、さらに、RAIDディスクを両方のノードに接続しておくことにより、片方のノードに障害が発生した場合でも、もう片方が処理を引き継いでサービスを継続して提供するフェイルオーバー機能を提供することができる。さらに、この技術は、各ノードのNVRAMに、それぞれ相手ノードのNFSコマンドログのコピーを格納するエリアを確保し、NFSコマンド受信時に、自ノードのNVRAM上にログを格納すると同時にネットワーク経由で相手ノードのNVRAMにもコピーを行っておくため、システム障害発生時に、自ノード上に保存されている相手ノードのNFSコマンドログをアクセスし、相手ノードが使用していたRAIDディスクのファイルシステムを復旧してサービスを継続することができ、高い可用性を提供することができる。

【0014】

【発明が解決しようとする課題】前述した従来技術は、SAN及びNASそれぞれの技術により管理コストを低減することができるが、SANへのストレージを提供するRAIDと、NASにおけるファイルサーバとを、それぞれ別の装置として実現しなければならないため、両方の機能が必要な場合、両方の装置を導入する必要があり、別の管理コストを増加させるという問題点があった。

【0015】また、前述の2ノード構成のNASにより高信頼性と高可用性とを提供することができる従来技術は、3ノード以上の構成のNASに適用する場合、各ノードのNVRAMにノード数分のログ格納領域を設ける必要があり、この結果、メモリ消費量が大きくなるという問題点を生じさせる。さらに、この従来技術は、NFSコマンドを受信する度に全ノードのNVRAMへのコピーを行う必要があり、ノード数の増加に応じて性能が低下するという問題を生じさせる。さらに、この従来技術は、複数のノードが同一のファイルシステムを変更する場合について配慮されていない。

【0016】本発明の目的は、前述した従来技術の問題点を解決し、SANインタフェースとNASインタフェースとの両方を提供し、障害発生時にもデータが失われることのない高信頼性と、かつ、任意の数のNASインタフェースが同一のファイルシステムに高性能アクセスすることを可能としたSAN/NASを統合したストレージシステム及びその制御方法を提供することにある。

【0017】また、本発明の他の目的は、同一のファイルシステムにアクセス可能なNASインタフェースの数

によらず、ファイルシステム変更ログ格納用のメモリサイズを一定にすることができ、かつ、そのサイズをユーザが指定可能なSAN/NASを統合したストレージシステム及びその制御方法を提供することにある。

【0018】さらに、本発明の他の目的は、あるNASインタフェースに障害が発生した場合にも、別のNASインタフェースがその処理を引き継ぐフェイルオーバー処理を可能にし、このフェイルオーバー処理を正常なNASインタフェースが存在する限り連続して実行することができる高可用性を持ったSAN/NASを統合したストレージシステム及びその制御方法を提供することにある。

【0019】

【課題を解決するための手段】本発明によれば前記目的は、外部ネットワーク接続用の複数のインタフェースと、該複数のインタフェースからアクセス可能な複数のディスクと、前記複数のインタフェースからアクセス可能な共有メモリとを備えたストレージシステムにおいて、前記複数のインタフェースには、ディスクブロック単位のI/O要求を処理するブロックインターフェースと、ファイル単位のI/O要求を処理するファイルサーバを搭載したファイルインタフェースとのいずれかあるいは両方が搭載され、前記複数のディスクの一部には、複数のファイルサーバが共有アクセス可能なファイルシステムが構築され、前記共有メモリには、前記ファイルシステムの変更ログを保持するログ格納領域と、前記ファイルシステムの排他制御及びログ格納領域の管理を行う管理用のファイルサーバに関連する情報を保持する管理ファイルサーバ情報格納領域とが構築されていることにより達成される。

【0020】また、前記目的は、外部ネットワーク接続用の複数のインタフェースと、該複数のインタフェースからアクセス可能な複数のディスクと、前記複数のインタフェースからアクセス可能な共有メモリとを備えたストレージシステムの制御方法において、前記複数のインタフェースには、ディスクブロック単位のI/O要求を処理するブロックインターフェースと、ファイル単位のI/O要求を処理するファイルサーバを搭載したファイルインタフェースとのいずれかあるいは両方が搭載され、前記複数のディスクの一部には、複数のファイルサーバが共有アクセス可能なファイルシステムが構築され、前記共有メモリには、前記ファイルシステムの変更ログを保持するログ格納領域と、前記ファイルシステムの排他制御及びログ格納領域の管理を行う管理用のファイルサーバに関連する情報を保持する管理ファイルサーバ情報格納領域とが構築されており、前記ファイルシステムの管理ファイルサーバ以外のファイルサーバが、外部ネットワークからファイル書き込み要求を受信して、そのファイル書き込み要求を解析して書き込み対象ファイルが含まれるファイルシステムの管理ファイルサーバ

を特定し、管理ファイルサーバに対してファイル書き込み情報を送信した後、その応答としてユーザデータを書き込むディスクブロック情報とログ格納領域内に割り当てられたログ格納アドレス情報とを受信し、受信したログ格納アドレス情報を用いてユーザデータ格納領域にユーザデータを格納した後、ログ格納領域内のログステータス情報を変更し、ディスクブロック情報に基づいて前記ディスクにユーザデータを格納した後、ログ格納領域内のログステータス情報を変更し、ファイル書き込み結果情報を前記ファイルシステムの管理ファイルサーバに送信した後、外部ネットワークを経由して受信したファイル書き込み要求の応答を、外部ネットワークに送信することにより達成される。

【0021】

【発明の実施の形態】以下、本発明によるストレージシステム及びその制御方法の実施形態を図面により詳細に説明する。

【0022】図1は本発明の一実施形態によるストレージシステムの構成を示すブロック図、図2はクライアントからのI/O要求を受信したファイルサーバの処理動作を説明するフローチャート、図3は管理ファイルサーバの処理動作を説明するフローチャート、図4は管理ファイルサーバによるファイルシステム復旧の処理動作を説明するフローチャート、図5は管理ファイルサーバ情報保持手段が持つ情報について説明する図である。図1において、100はストレージシステム、110、120、130はファイルインタフェースボード、111はFS（ファイルシステム）管理情報保持手段、112はログ格納領域管理手段、113、123、132はファイルサーバ、122はデータ一時蓄積部、140はiSCSIインタフェースボード、150はFC/SCSIインタフェースボード、160、170はディスク、172はファイルシステム、180は共有メモリ、181は交代サーバ決定手段、182は管理ファイルサーバ情報保持手段、183はファイルサーバ状態保持テーブル、186はログ格納領域、190は内部ネットワーク、200、300、400、500はクライアントホスト、600はSAN、700はLANである。

【0023】本発明の実施形態によるストレージシステム100は、LAN700及びSAN600に接続され、LAN700に接続されるクライアントホスト200、300、400、及び、SAN600に接続されるクライアントホスト400、500からネットワークを経由して受信するI/O要求を処理する。このストレージシステム100は、ファイルサーバ113、123、132を搭載したファイルインタフェースボード110、120、130と、iSCSIコマンドを処理するiSCSIインタフェースボード140と、ファイバチャネルを介して受信したSCSIコマンドを処理するFC/SCSIインタフェースボード150と、ディス



ク160、170及び共有メモリ180とを備え、前述の各ボードとディスク160、170及び共有メモリ180とが内部ネットワーク190で結合されて構成されている。

【0024】ストレージシステム100は、前述したインタフェースボードをそれぞれ異なるスロットに装着している。なお、各スロットは、任意のインタフェースボードをシステム稼動時に挿抜可能である。従って、本発明の実施形態によるストレージシステム100は、ユーザの用途に応じて各インタフェースボードの割合を動的に変更することができる。

【0025】ディスク160、170は、前述の各インタフェースボードのいずれからもアクセス可能である。そして、ディスク160は、iSCSIインタフェースボード140及びFC/SCSIインタフェースボード150がブロック単位のアクセスを行うためのディスクであり、ディスク170には、ファイルサーバがアクセスするためのファイルシステム172が構築されている。なお、ここでのディスクは、RAIDを含む論理的なディスクを意味しており、簡単のためにそれぞれ1つつつ記載しているが、実際には任意の数のディスクを搭載することが可能である。

【0026】ファイルサーバ113は、ファイルシステム172の管理ファイルサーバであり、他のファイルサーバがファイルシステム172をアクセスする場合に、必ずファイルサーバ113と通信を行う。ファイルサーバ113は、ファイルシステム管理情報保持手段111とログ管理手段112とを有している。ファイルシステム管理情報保持手段111は、ファイルの属性やファイルが保持するディスクブロックのリスト、ファイルシステム内の未使用ディスクブロックリスト等のメタデータや、ファイルデータを保持しており、ファイルシステムへの変更が発生するたびに内部の情報を更新する。但し、これらのデータは、ディスク170上で複数のディスクブロックにまたがって格納されるため、変更の度にディスクへ格納すると性能が大きく低下する。このため、ディスクへのデータの書き戻しは、クライアントホストからの書き戻し要求の発生時か、または、ある一定時間毎に実行する。

【0027】変更データのディスクへの反映をファイルシステム変更処理と非同期に行うと、ファイルサーバに障害が発生した場合に、ディスクに未反映となっている変更データが失われる危険性がある。本発明の実施形態は、この問題を防止するため、ファイルシステムの変更データのログを、不揮発性の共有メモリ180上に設けたログ格納領域186に格納する。ログとして格納するログエントリは、変更メタデータ184と変更ファイルデータ185とにより構成される。ファイルサーバ113は、ログ管理手段112を用いてログ格納領域186を環状に使用し、ファイルシステムの変更データをディ

スクに反映した場合に、対応する変更データを含むログエントリを解放する。

【0028】本発明の実施形態によるストレージシステムは、前述のように、ファイルシステムの変更と同期的にログを格納することとしているので、全ての変更情報を残すことができ、また、共有メモリ180が不揮発性であるため、変更データが失われることがない。また、ファイルサーバ113に障害が発生した場合にも、再起動後にログ格納領域186内のログデータを参照しファイルシステムを最新の状態に修復することが可能となる。

【0029】ファイルサーバ123は、ファイルシステム172の管理ファイルサーバではないが、クライアントホストからファイルシステム172上のファイルに対するI/O要求を受信した場合に、ファイルシステム172をアクセスする。ファイルサーバ123は、ファイルシステム172のメタデータ情報を持たないため、ファイルシステム172の管理ファイルサーバであるファイルサーバ113を特定し、必要な処理を依頼する必要がある。I/O要求がメタデータアクセスのみの場合、全てのファイルシステム処理は、ファイルサーバ113で完結する。

【0030】一方、READやWRITEといったファイルデータへのアクセスが必要な処理の場合、メタデータの処理は、ファイルサーバ113で行われ、ファイルデータの転送処理は、ファイルサーバ123で行われる。これにより、管理ファイルサーバの負荷を低減すると共に、ファイルデータのファイルサーバ間コピーを不要とすることができる。

【0031】図5に示すように、管理サーバ情報保持手段182は、ログ格納領域管理テーブル5000とディスク管理テーブル5100とにより構成される。ログ格納領域管理テーブル5000は、管理ファイルサーバの番号5001と、そのログ格納領域アドレス5002と、ログ格納領域サイズ5003とからなる。管理ファイルサーバ113は、ファイルシステム172を正常にマウントするとき、ログ格納領域を割り当て、その先頭アドレス及びサイズをこのテーブル5000に格納し、異常終了後にファイルシステムをマウントするとき、このテーブル5000を参照してファイルシステムの復旧を行う。なお、複数のファイルシステムが備えられる場合に、複数のファイルシステムのそれぞれに対するファイル管理サーバは、異なったものであってもよい。

【0032】ディスク管理テーブル5100は、各ディスクのディスク番号5001、デフォルト管理ファイルサーバ5102、カレント管理ファイルサーバ5103、ログ種別5103を保持している。ストレージシステム100は、管理ファイルサーバのフェイルオーバー機能を提供しており、デフォルト管理ファイルサーバ5102は、フェイルオーバーが発生していない場合にそのデ



ィスクを管理すべき管理ファイルサーバが登録される。カレント管理ファイルサーバ5103は、フェイルオーバーの有無にかかわらず現在そのディスクを管理している管理ファイルサーバが登録される。

【0033】ログ種別5104は、ログに格納する変更データの種別を記述する。変更データの種別としては、変更メタデータ、ファイルデータの2種類があり、ログ種別5104にはその両方を格納するか、変更メタデータのみを格納するか、あるいは、全くログを採取しないかの3つのうちの1つが登録される。変更メタデータとファイルデータとの両方をログに格納すれば、全てのデータを完全に保証することができる。変更メタデータのみをログに格納する場合、ファイルシステムの整合性を保証することができるが、ファイルデータが失われる可能性がある。また、全くログを採取しない場合、ファイルシステムの整合性もファイルデータも失われる可能性がある。この場合システム障害発生後、最初にファイルシステムを使用する前に、ファイルシステムの整合性チェックの処理と修復を行うfsckのような処理とを行う必要がある。一般に、このfsckの処理時間は、ファイルシステムサイズに比例して増加し、数分から数十分を要することがあるため、フェイルオーバー時には使用できない。データの信頼性保証と性能とはトレードオフの関係にあるため、信頼性を向上させるとファイルシステムアクセス性能は低下する。このため、本発明の実施形態では、ログ種別5104は、ユーザが用途に応じて指定することができるようにしている。

【0034】次に、図2に示すフローを参照して、ファイルサーバ123がクライアントホスト300からファイルシステム172に対するI/O要求を受信した場合の処理動作を説明する。

【0035】(1) まず、ファイルサーバ123は、クライアントホスト300から受信したI/O要求を解析し、ファイルシステム172へのI/O要求であることを検出すると、ディスク管理テーブル5100を検索し、ディスク番号5101がファイルシステム172のディスク番号と一致するエントリを特定し、そのカレント管理ファイルサーバ5103に格納されている管理ファイルサーバを読み出す。ここでは、管理ファイルサーバとして、ファイルサーバ113が設定されているものとする(ステップ2001)。

【0036】(2) 次に、管理ファイルサーバ113にファイルアクセス要求を送信する。このとき、I/O要求がWRITE要求の場合、要求に伴って受信しているファイルデータは、ファイルサーバ123内のデータ一時蓄積部122に残し、ファイルサーバ113には送信しない(ステップ2002(図1の■))。

【0037】(3) 次に、I/O要求がREADやWRITE等のファイルデータをアクセスする要求か否かを判断する。ファイルデータをアクセスしないI/O要求

の場合、その処理は、管理ファイルサーバ113のみで完結するので、その場合、ファイルサーバ123は、ファイルサーバ113からの処理結果を受信した後、その処理結果をクライアントホスト300に返信して処理を終了する(ステップ2003、2014、2015)。

【0038】(4) ステップ2003の判定で、I/O要求がファイルデータをアクセスする要求の場合、ファイルサーバ113からファイルデータを保存するディスクブロック番号と、ログ格納アドレス情報とを受信する。このとき、ログ格納アドレス情報の受信は、WRITE要求の場合のみであり、READ要求のときには不要である(ステップ2004(図1の■))。

【0039】(5) 次に、I/O要求がWRITE要求であるか否かを判断する。I/O要求がWRITEでな居場合、すなわち、READの場合、ファイルデータの格納されているディスクブロック番号のみをファイル管理サーバ113から受信しているため、受信したディスクブロック番号を使用してディスク170をアクセスしてファイルデータを読み出し、読み出したファイルデータをクライアントホスト300に返信し、その後、処理結果をファイルサーバ113に返信して処理を終了する(ステップ2005、2012、2013、2011)。

【0040】(6) ステップ2005の判定で、I/O要求がWRITEであった場合、ファイルサーバ113から受信したログ格納アドレス情報を用い、ファイルデータをデータ一時蓄積部122から変更ファイルデータ格納領域185に格納する。その後、変更メタデータ格納領域184に格納されているログステータス情報を“データログ書き込み”の状態に変更する(ステップ2006(図1の■)、2007(図1の■))。

【0041】(7) ここまでの処理により、クライアント300からのI/O要求の全てがログ格納領域186に反映されたため、以後、ファイルサーバ113に障害が発生しても、ログ情報を用いてファイルシステム172を復旧することができる。このため、ファイルサーバ123は、クライアントホスト300に対して処理結果を返信する(ステップ2008)。

【0042】(8) 次に、ファイルサーバ123は、ファイルサーバ113から受信したディスクブロック情報を用いて、データ一時蓄積部122に格納していたファイルデータをディスク170に格納し、その後、ログステータス情報を“データディスク書き込み”状態に変更し、さらに、ファイルサーバ113に処理結果を返信して処理を終了する(ステップ2009(図1の■)~2011(図1の■))。

【0043】ファイルサーバ123によるI/O要求に対する処理は、前述した通りであるが、ファイルサーバ123以外にファイルシステム172をアクセスするファイルサーバが存在した場合にも、ファイルデータのロ

格納領域 186 への格納及びディスク 170 へのアクセス処理は、それぞれ並列に実行することができるため高速な I/O 処理性能を提供できる。

【0044】次に、図 3 に示すフローを参照して、管理ファイルサーバ 113 での処理動作を説明する。ここでは、ファイルサーバ 123 からのファイルアクセス要求を受信した場合の処理について説明するが、その他のファイルサーバからファイルアクセス要求を受信した場合も同様である。

【0045】(1) 管理ファイルサーバ 113 は、ファイルサーバ 123 からのファイルアクセス要求を受信すると、そのファイルアクセス要求を解析し、ファイルシステム管理情報保持手段 111 を用いてアクセスするファイルを特定すると同時に、ファイルに対してロックをかける (ステップ 3001)。

【0046】(2) 次に、ファイルアクセス要求が READ や WRITE 等のファイルデータアクセス要求か、それ以外の要求かを判断する。要求がファイルデータアクセス要求以外の場合、ファイルサーバ 123 にディスクブロック番号やログ格納アドレス情報を返信する必要がなく、自サーバによるメタデータアクセスのみで処理が終了する。このため、管理ファイルサーバ 113 は、メタデータアクセスを行い、メタデータの変更があった場合に、ログ管理手段 112 が割り当てたログエントリに変更メタデータを格納する。その後、ファイルのロックを解除し、処理結果をファイルサーバ 123 に返信して処理を終了する (ステップ 3002、3009～3011)。

【0047】(3) ステップ 3002 の判定で、要求がファイルデータアクセス要求であった場合、次に、I/O 要求が WRITE 要求であるか否かを判断する。WRITE 要求でなかった場合、すなわち、READ 要求であった場合、ファイルシステムに対する変更が発生しないためログ採取の必要がない。従って、管理ファイルサーバ 113 は、メタデータをアクセスし、ファイルデータが格納されているディスクブロック番号を算出し、ファイルサーバ 123 にディスクブロック番号を返信する (ステップ 3003、3012、3013)。

【0048】(4) ステップ 3003 の判定で、I/O 要求が WRITE 要求であった場合、ファイルデータを格納するディスクブロックを割り当てると共に、メタデータの変更を行う。次に、変更メタデータと WRITE するファイルデータを格納するログエントリサイズを算出し、ログ管理手段 112 を用いて、ログ格納領域 186 内に変更メタデータ格納領域 184 とファイルデータ格納領域 185 とを含むログエントリを割り当てる (ステップ 3004、3005)。

【0049】(5) その後、変更メタデータ内のログステータス情報を“データ未書き込み”に設定した後、変更メタデータを変更メタデータ格納領域 184 に格納

し、次に、変更メタデータ格納領域 184 及びファイルデータ格納領域 185 のアドレス及びサイズと、ステップ 3004 の処理で割り当てたディスクブロック番号をファイルサーバ 123 に返信する (ステップ 3006

50 (図 1 の■)、3007 (図 1 の■))。

【0050】(6) ステップ 3007 またはステップ 3013 の処理後、ファイルサーバ 123 からのファイルアクセス終了のメッセージを受信すると、ファイルのロックを解除して処理を終了する (ステップ 3008)。

10 【0051】次に、図 4 に示すフローを参照して、管理ファイルサーバ 113 がファイルシステム 172 を復旧する処理動作を説明する。ここでは、管理ファイルサーバ 113 に障害が発生し、その後、管理ファイルサーバ 113 が障害から回復し、ファイルシステム 172 を復旧するものとしている。

15 【0052】(1) 管理ファイルサーバ 113 は、まず、ファイルシステム 172 を格納されているディスク番号を用いてディスク管理テーブル 5100 を検索し、ディスク 170 を格納しているエントリを特定し、そのカレント管理ファイルサーバフィールド 5103 を参照して、直前にファイルシステム 172 を管理していた管理ファイルサーバを特定する。ここでは、ファイルサーバ 113 が格納されているものとする。さらに、ログ種別フィールド 5104 を参照して、ログ種別を取得する (ステップ 4001)。

20 【0053】(2) 次に、ステップ 4002 で取得したログ種別が“ログ不採取”であるか否かを判断し、“ログ不採取”の場合、ログを用いた復旧が不可能であるため、ファイルシステムチェックプログラムである fsck を起動し、ファイルシステムの修復を行った後、ファイルシステム復旧処理を終了する (ステップ 4002、4010)。

25 【0054】(3) ステップ 4002 の判定で、ログ種別が“ログ不採取”でなかった場合、ログ格納領域管理テーブル 5000 を検索し、管理ファイルサーバ 113 のログ格納領域のアドレス及びサイズを求める。その後、ログ格納領域 186 をスキャンして、ログエントリのポインタをディスク 170 に未反映の変更ログを保持しているログエントリに設定する (ステップ 4003、4004)。

35 【0055】(4) 次に、ファイルシステム 172 に対する全ての変更ログをディスク 170 に反映したか否かをチェックし、全て反映済みであれば、ファイルシステム復旧処理を終了する。また、まだ未反映のログエントリが存在する場合、ポインタがさしているログエントリを参照し、そのログステータス情報が“データログ書き込み”状態か否かを調べる (ステップ 4005、4006)。

45 【0056】(5) ステップ 4006 の判定で、ログステータス情報が“データログ書き込み”状態であった場

合、ファイルデータがログに格納されているがディスクには未反映の状態を示しているため、ログに格納されているファイルデータを、メタデータ変更ログのディスクブロック情報に従ってディスクに格納する（ステップ4007）。

【0057】（6）ステップ4007の処理後、変更メタデータをファイルシステムの管理自に反映させ、ファイルシステム172に関する次の未反映ログエントリにポインタを進める。そして、ステップ4005の処理に戻って、ここからの処理を繰り返す（ステップ4008、4009）。

【0058】（7）また、ステップ4006の判定で、ログステータス情報が“データログ書き込み”状態ではなかった場合、ファイルデータ格納ログ領域内のデータは意味が無いため、ステップ4007の処理を実行することなく、ステップ4008からの処理を実行する。

【0059】図4により説明した処理を、全てのログエントリについて繰り返して実行することにより、ファイルシステムを最新の状態に復旧することができる。ログ格納領域186に格納されている全てのログエントリの処理が終了すると、ログ格納領域186内の全てのログエントリは解放される。

【0060】前述したような構成を持ち、前述したような処理を行う本発明の実施形態によれば、SANインタフェースとNASインタフェースとの両方を任意の割合で提供し、障害発生時にもデータを失わない高信頼性を得ることができ、かつ、任意の数のNASインタフェースが同一のファイルシステムに高性能なアクセスをすることを可能としたSAN/NAS統合ストレージシステムを提供することができる。

【0061】図6はログ格納領域のサイズをクライアントホストから設定する際の管理ファイルサーバの処理動作の例を説明するフローチャートであり、次に、これについて説明する。

【0062】（1）管理ファイルサーバ113は、クライアントホストからのログ格納領域サイズ設定コマンドを受信すると、受信したサイズの値が正常範囲内であるか否かを判定し、サイズが正常範囲外であれば、処理結果としてエラーをクライアントホストに返信して、ここでの処理を終了する（ステップ6001、6002、6008）。

【0063】（2）ステップ6002の判定で、ログ格納領域サイズが正常範囲内であることが確認できると、ログ格納領域186が使用中（領域内にデータがある）か否かをログ管理手段112により調べて判定する。ログ格納領域が使用中の場合、次に、新たなI/O要求処理の開始を抑止すると共に、現在実行中のI/O処理終了を待ち合わせる（ステップ6003、6004）。

【0064】（3）その後、ステップ6005でファイルシステム管理情報保持手段111内にあるディスク未

反映データを全てディスクに反映させる。その結果としてログ格納領域186内の全てのログエントリは解放され、ログ格納領域186が未使用の状態と等しくなる（ステップ6005）。

05 【0065】（4）ステップ6005の処理のご、または、ステップ6003での判定で、ログ格納領域186が使用中でなかった場合、指定されたサイズのログ格納領域を共有メモリ180内に確保し、そのアドレスとサイズとをログ格納領域管理テーブル5000内のログ格納領域アドレスフィールド5001とログ格納領域サイズフィールド5002とに格納する。その後、I/O処理を再開し、処理結果をクライアントに返信して処理を終了する（ステップ6006～6008）。

15 【0066】前述した図6に示す例は、ログ格納領域サイズ変更のコマンドを送信したクライアントホストが適切な権限をもっているか否かのチェックを行っていないが、コマンド受信後に適切な権限チェックを行うようにすることもできる。

20 【0067】本発明の実施形態は、前述した処理を行うことにより、同一のファイルシステムにアクセス可能なNASインタフェースの数によらず、ファイルシステム変更ログ格納用のメモリサイズを一定にすることができ、かつ、そのサイズをクライアントが指定した値に設定することが可能となる。

25 【0068】図7はファイルサーバ状態テーブル183の構成を説明する図、図8はファイルサーバ132によるファイルサーバ113の監視及びフェイルオーバーの処理動作を説明するフローチャートであり、次に、管理ファイルサーバに障害が発生した場合に他のファイルサーバが管理ファイルサーバの処理を引き継ぐフェイルオーバー処理について説明する。

30 【0069】ここで説明する例は、図1において、共有メモリ180内にある交替サーバ決定手段181が、ある管理ファイルサーバの状態を監視し、異常発生時にはその処理を引き継ぐべき交代ファイルサーバを決定し、交代サーバ決定手段181によって決定された交代ファイルサーバが、ファイルサーバ状態テーブル183を用いて、監視対象の管理ファイルサーバの状態を常に監視し、異常を検知すると即座にフェイルオーバー処理を開始する処理である。説明する例では、管理ファイルサーバ113の交代ファイルサーバがファイルサーバ132であると

45 【0070】ファイルサーバ状態テーブル183は、図7に示すように、ファイルサーバ番号7001、状態7002、タイムスタンプ7003、IPアドレスやMACアドレスなどのネットワーク情報7004から構成される。各ファイルサーバは、一定のリフレッシュ時間間隔で、ファイルサーバ状態テーブル183のタイムスタンプ7003を更新する。交代ファイルサーバは、監視対象のファイルサーバのタイムスタンプ7003を一定

のチェック時間間隔で検査し、タイムスタンプ7003の値が正しく更新されていれば、監視対象のファイルサーバが正しく動作していると判断し、タイムスタンプ7003の値が更新されていなければ異常が発生したと判断してフェイルオーバー処理を開始する。ここで、チェック時間間隔はリフレッシュ時間間隔よりも大きな値である必要がある。

【0071】次に、図8に示すフローを参照して、ファイルサーバ132によるファイルサーバ113の監視及びフェイルオーバー処理の動作を説明する。

【0072】(1) ファイルサーバ132は、チェック時間になるとファイルサーバ状態テーブル183から監視対象であるファイルサーバ113のエントリを検索して、そのタイムスタンプ7003をチェックし、タイムスタンプが正しく更新されているかを判定する。タイムスタンプが正しく更新されている場合、何も行わずに処理を終了する(ステップ8001、8002)。

【0073】(2) ステップ8002の判定で、タイムスタンプが正しく更新されていないことを検知すると、ファイルサーバ132は、フェイルオーバー処理を開始する。そして、まず、ファイルサーバ状態テーブル183のファイルサーバ113の状態7001を“フェイルオーバー処理中”に設定して2重フェイルオーバー処理の起動を抑止する(ステップ8003)。

【0074】(3) 次に、ディスク管理テーブル5100を検索し、デフォルト管理ファイルサーバ5102がファイルサーバ113であるディスクのエントリを取得し、取得したディスクの復旧処理を行う。ここでの復旧処理は、図4により説明した復旧処理と同様に行われる(ステップ8004、8005)。

【0075】(4) ディスクの復旧処理が終了すると、次に、ディスク管理テーブルのカレントファイルサーバに交代ファイルサーバであるファイルサーバ132の番号を格納し、ファイルサーバ113が管理していた全ディスクの復旧が終了したか否かを判断する。まだ、復旧処理の終了していないディスクが存在する場合、ステップ8004の処理に戻って以後の処理を繰り返す(ステップ8006、8007)。

【0076】(5) ファイルサーバ113が管理していた全てのディスクに関する復旧処理が終了すると、次に、ファイルサーバ状態テーブルのネットワーク情報7004を参照し、ファイルサーバ113のネットワークアダプタに設定されていた情報をファイルサーバ132のネットワークアダプタに引き継ぐ(ステップ8008)。

【0077】(6) その後、ファイルサーバ状態テーブルにあるファイルサーバ113の状態7002を“フェイルオーバー”状態に変更し、最後に、ファイルサーバ132がファイル管理サーバ変更通知手段131を用いて他のファイルサーバに管理ファイルサーバの変更を通知

してフェイルオーバー処理を終了する(ステップ8009、8010)。

【0078】前述の処理において、管理ファイルサーバ変更通知を受け取ったファイルサーバ123は、管理ファイルサーバ情報保持手段182の情報を再び共有メモリ180から読み込み、ファイルサーバ123内の管理サーバ情報保持手段121に格納する。また、交代サーバ決定手段181は、交代ファイルサーバ132がファイルサーバ113の異常を検知した時点で、全ファイルサーバの交代ファイルサーバを再決定し、各ファイルサーバに通知する。この処理により、常に正常な状態のファイルサーバ同士がお互いに監視し合うことが可能となる。

【0079】本発明の実施形態は、図8により説明した処理を行うことにより、あるNASインタフェースに障害が発生しても、他に正常なNASインタフェースが存在する限り、フェイルオーバー処理を用いてサービスを継続でき、高可用性を図ることができる。

【0080】  
【発明の効果】以上説明したように本発明によれば、複数のNASインタフェースが同一のファイルシステムをアクセスすることができるため、インタフェース数に比例した性能を提供しながら、障害発生時にもデータが失われることのない高い信頼性を得ることができ、さらに、正常なNASインタフェースが1つでも存在する限り継続してファイルアクセスサービスを行うことができる。

#### 【図面の簡単な説明】

【図1】本発明の一実施形態によるストレージシステムの構成を示すブロック図である。

【図2】クライアントからのI/O要求を受信したファイルサーバの処理動作を説明するフローチャートである。

【図3】管理ファイルサーバの処理動作を説明するフローチャートである。

【図4】管理ファイルサーバによるファイルシステム復旧の処理動作を説明するフローチャートである。

【図5】管理ファイルサーバ情報保持手段が持つ情報について説明する図である。

【図6】ログ格納領域のサイズをクライアントホストから設定する際の管理ファイルサーバの処理動作の例を説明するフローチャートである。

【図7】ファイルサーバ状態テーブルの構成を説明する図である。

【図8】ファイルサーバによる管理ファイルサーバの監視及びフェイルオーバーの処理動作を説明するフローチャートである。

#### 【符号の説明】

100 ストレージシステム

110、120、130 ファイルインタフェースボー

ド

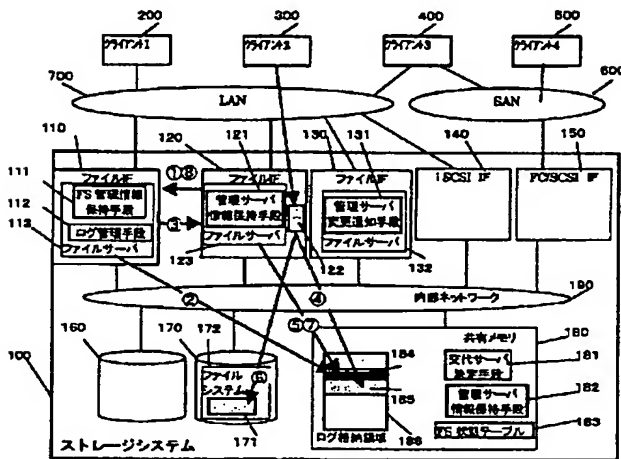
- 111 FS (ファイルシステム) 管理情報保持手段
- 112 ログ格納領域管理手段
- 113、123、132 ファイルサーバ
- 122 データ一時蓄積部
- 140 iSCSI インタフェースボード
- 150 FC/SCSI インタフェースボード
- 160、170 ディスク
- 172 ファイルシステム

05

- 180 共有メモリ
- 181 交代サーバ決定手段
- 182 管理ファイルサーバ情報保持手段
- 183 ファイルサーバ状態保持テーブル
- 186 ログ格納領域
- 190 内部ネットワーク
- 200、300、400、500 クライアントホスト
- 600 SAN
- 700 LAN

【図1】

図1



【図7】

図7

183

ファイルサーバ状態テーブル			
7001	7002	7003	7004
管理ファイルサーバ番号	状態	タイムスタンプ	ネットワーク情報

【図5】

図5

182

5000

ログ格納領域管理テーブル

5001	5002	5003
管理ファイルサーバ番号	ログ格納領域アドレス	ログ格納領域サイズ

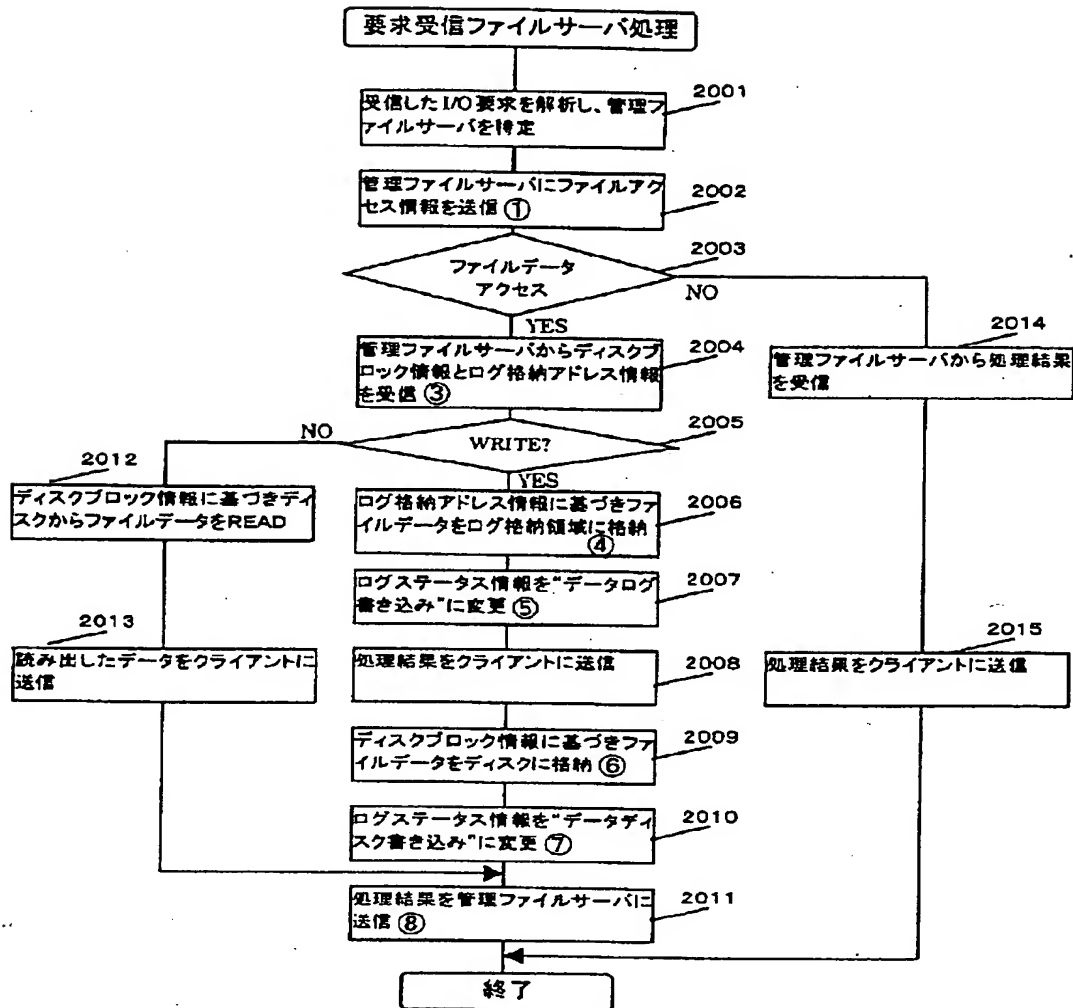
5100

ディスク管理テーブル

5101	5102	5103	5104
ディスク番号	デフォルト管理ファイルサーバ	カレント管理ファイルサーバ	ログ種別

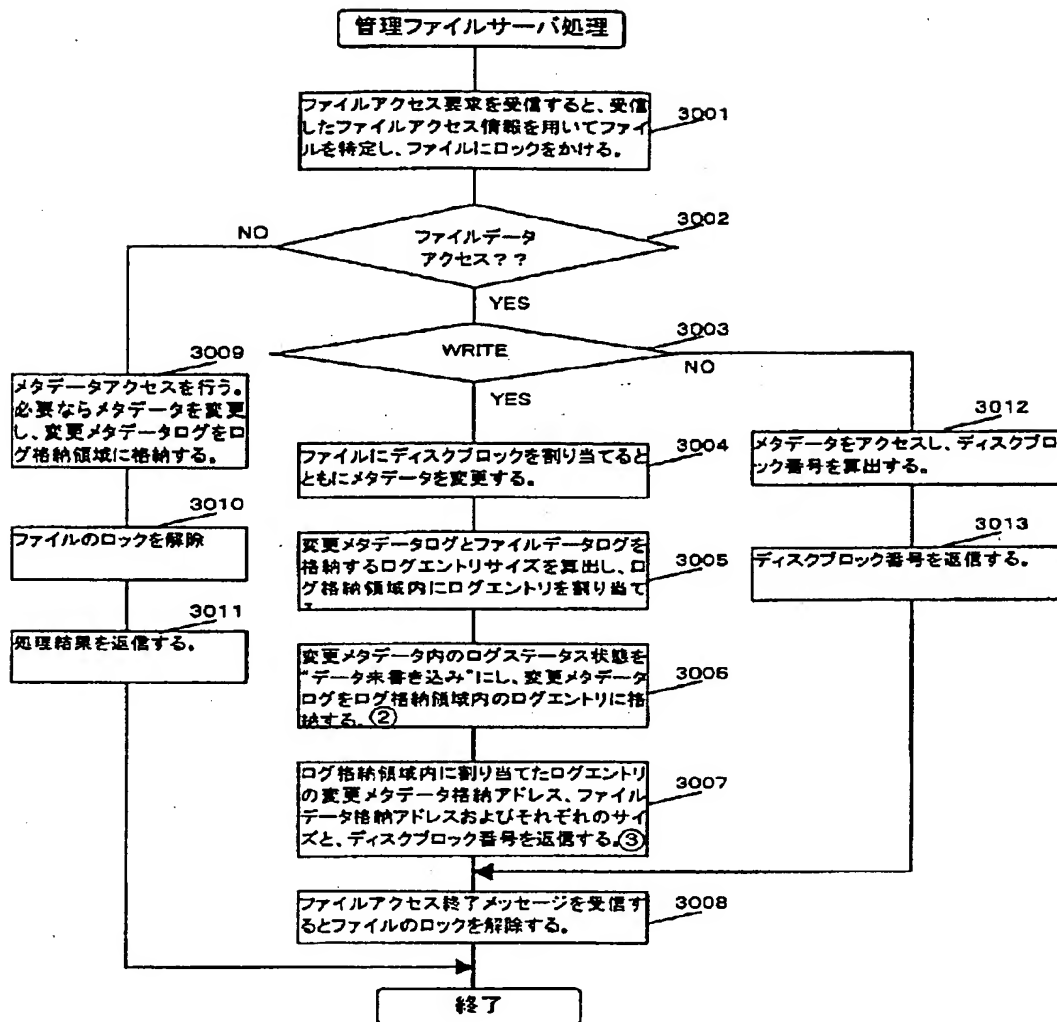
【図2】

図2



【図3】

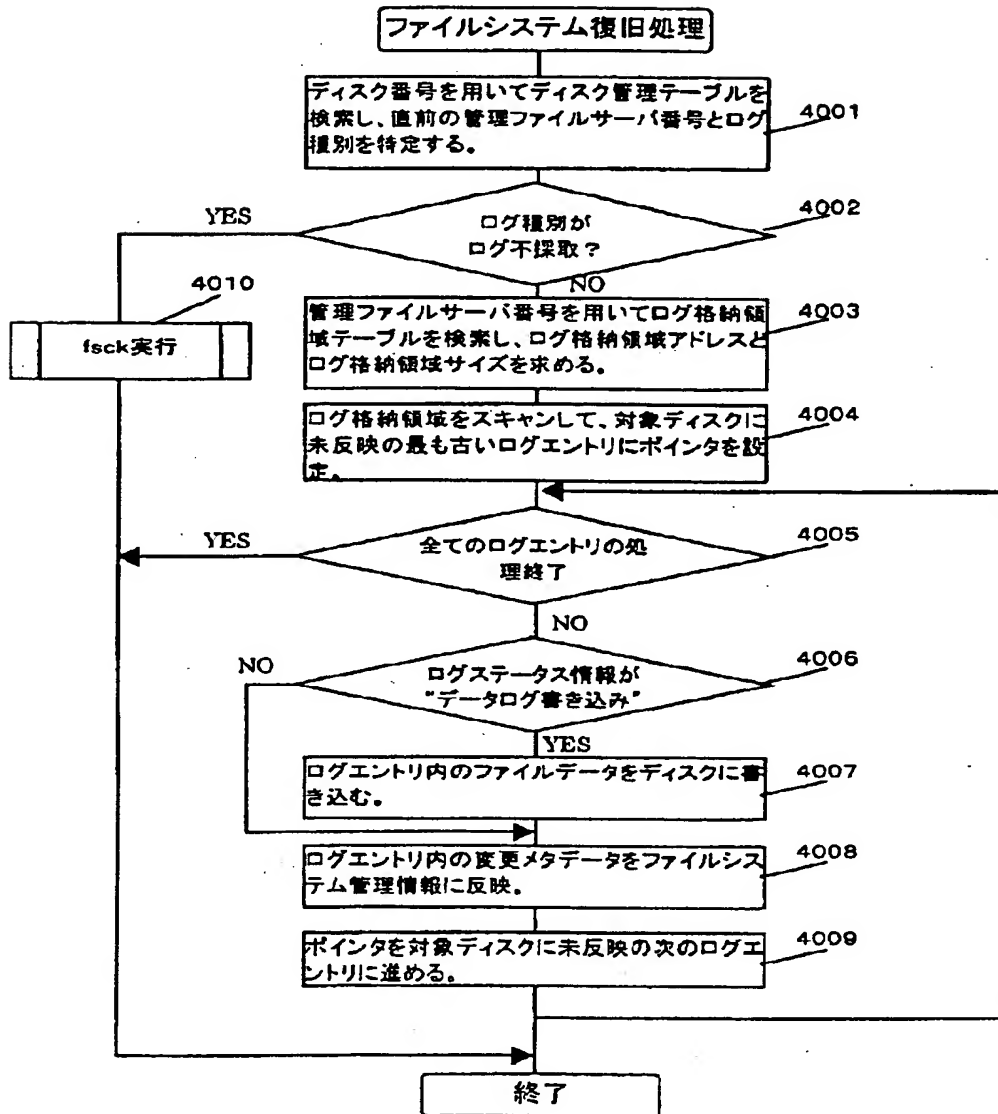
図3





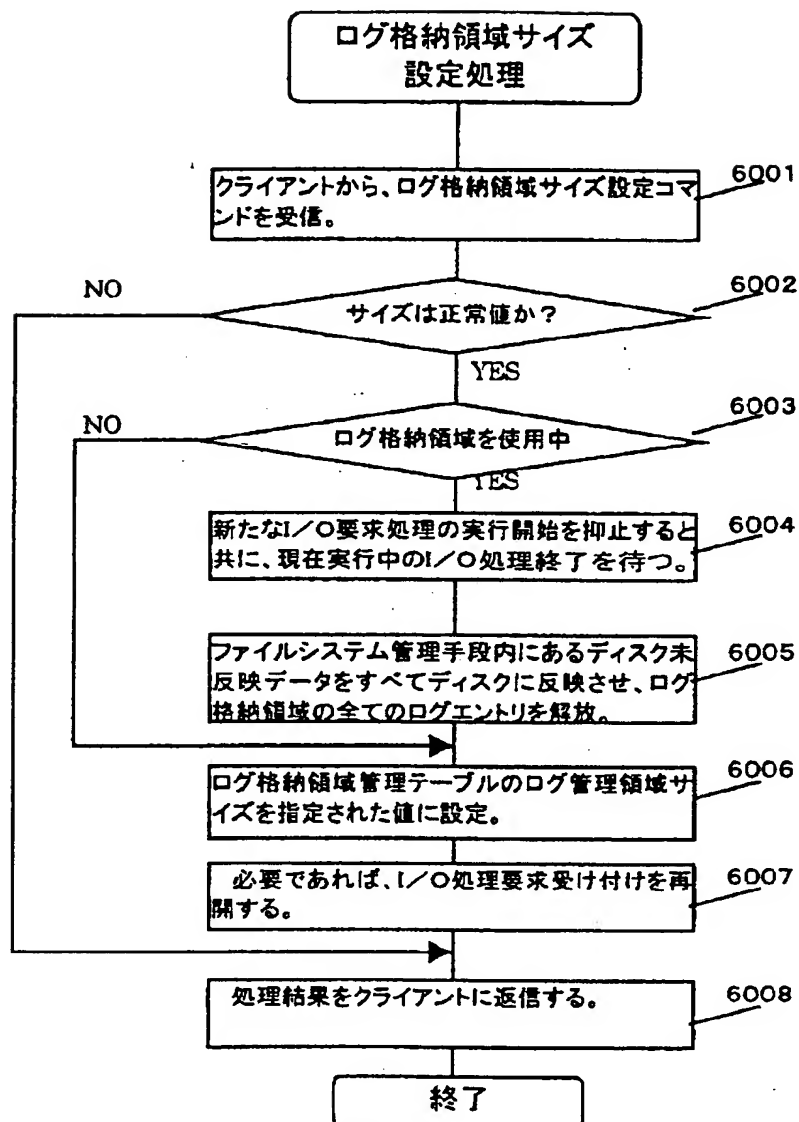
【図4】

図4



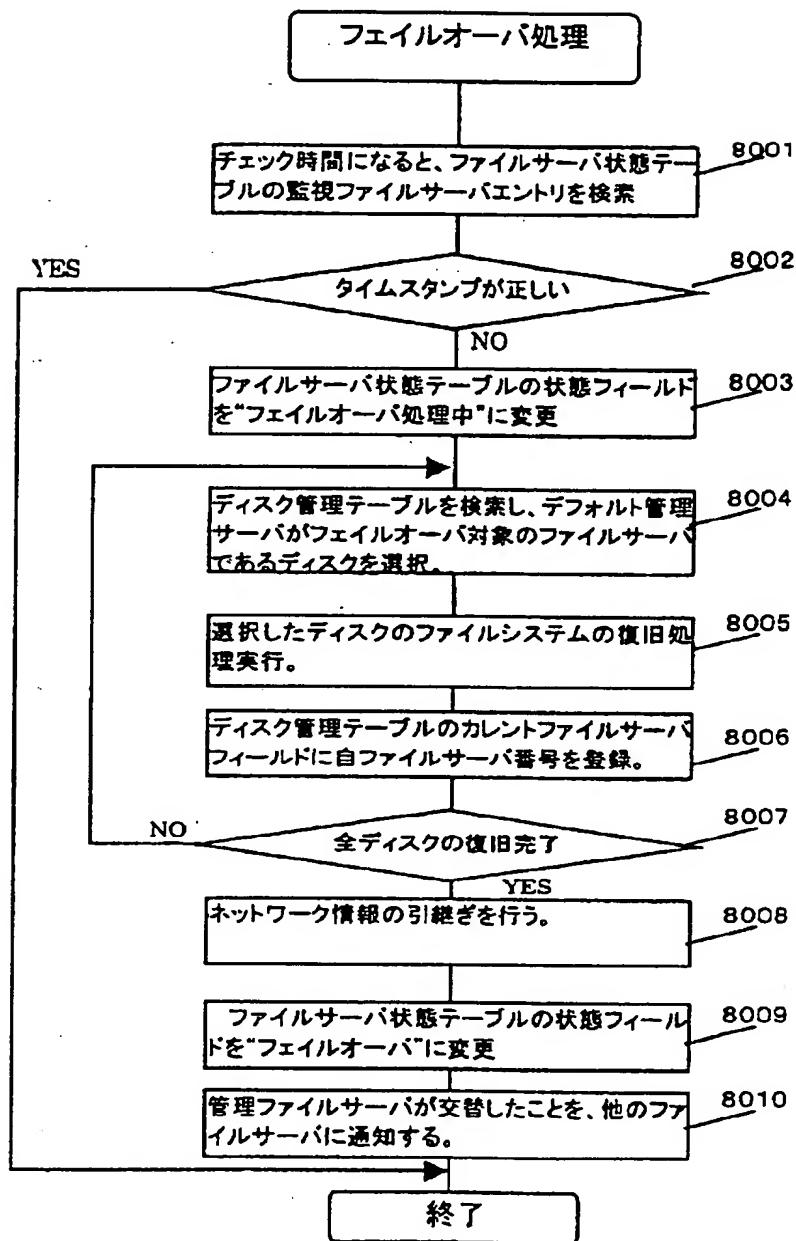
【図 6】

图6



【図8】

図8



フロントページの続き

(51) Int. Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)
G 0 6 F 3/06	5 4 0	G 0 6 F 3/06	5 4 0
(72) 発明者 北村 学		(72) 発明者 高田 豊	
神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内		神奈川県小田原市中里322番地2号 株式会社日立製作所RAIDシステム事業部内	
		Fターム(参考) 5B065 BA01 CA18 ZA01 ZA15 ZA17	
		5B082 DD08 FA17 GB02 GB06 HA08	

10